Redes Neuronales y Aprendizaje profundo

Tema 4 – Redes Generativas y Estado del Arte

Irina Arévalo

Contenido

1. Redes generativas
2. Aprendizaje por refuerzo
3. Sistemas de recomendación
4. Federated Learning

1. Redes Generativas

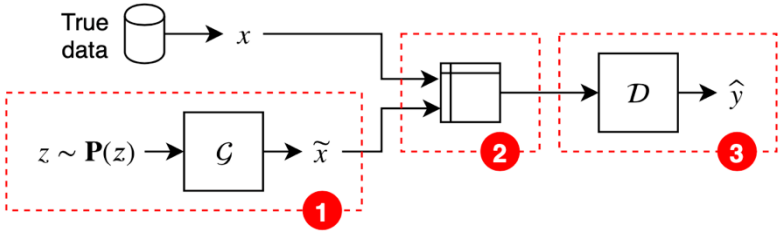
- Ya vimos que los sistemas de encoders/decoders se pueden usar para obtener un modelo generativo.

- Sin embargo, el modelo más habitual para esta tarea son las redes generativas adversarias (Generative Adversarial Networks, GAN)

- En las GANs encontramos dos redes con intereses contrapuestos, por eso se denominan adversarias. La red Generadora se entrena para producir datos (típicamente imágenes) falsos pero con tal grado de verosimilitud que sean capaces de confundir a una segunda red Discriminadora. Esta segunda red, a su vez, se entrena para distinguir los datos verdaderos de los falsos, generados por la otra otra red.

1. Redes Generativas

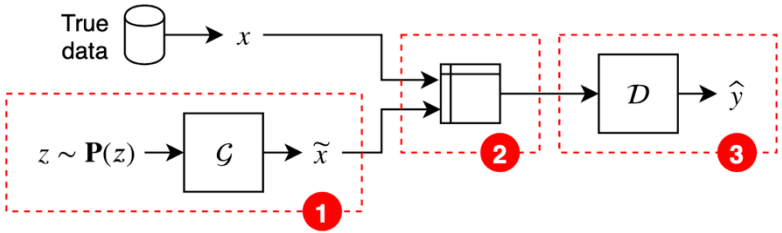
- Una red GAN se compone de dos redes neuronales con propósitos contrarios: la red generativa G y la red discriminativa D.
- La red generativa G recibe un vector *d*-dimensional *z* muestreado de una distribución de probabilidad *P(z),* que habitualmente es la Uniforme($[0,1]^d$). Este vector *z* hace las veces de una representación en el espacio latente de los datos reales, *x*. La red generativa transforma *z* en $\tilde{x} = G(z)$, que habitualmente se conoce como datos "falsos" (fake) (recuadro 1).
- El objetivo es lograr que la red generativa produzca datos falsos que puedan "hacerse pasar por verdaderos". En una primera aproximación podríamos pensar en minimizar la diferencia entre $\tilde{x}$ y *x*. Pero de este modo lograríamos aprender a reproducir los datos verdaderos, no a producir nuevas muestras de los datos.



1. Redes Generativas

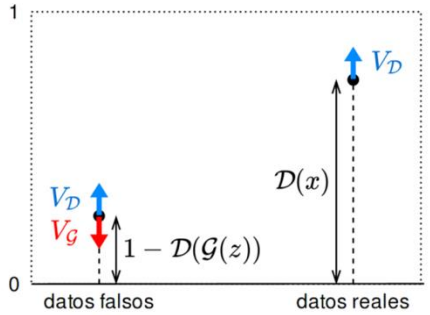
- La solución que proporcionan las GAN consiste en crear un nuevo conjunto de datos juntando muestras verdaderas, *x*, con muestras falsas, $\tilde{x}$. De este modo logramos recopilar un conjunto etiquetado de datos de un mismo tipo ya que sabemos cuales son verdaderos (*y = 1*) y cuales falsos (*y = 0*) (parte 2).

- La red discriminativa D recibe este segundo conjunto de datos y se entrena para producir la probabilidad de que el ejemplo introducido sea verdadero. Para ello la capa de salida de D consta de una única neurona con activación sigmoide (parte 3)



1. Redes Generativas

- El proceso de entrenamiento comienza con la primera etapa forward para producir la salida $\hat{y}$. En la segunda etapa se ajustan los pesos de las redes *D* y *G*. Para ello necesitamos una función de pérdida que será el motor del algoritmo de back-propagation. Pero primero hay que formalizar el objetivo de cada una de las redes.
- Por un lado, el objetivo de la red *D* es maximizar el valor de su salida $\hat{y}$ cuando la entrada es un dato real *x*, y al mismo tiempo minimizar el valor de $\hat{y}$ cuando la entrada es un dato falso; pero esto es equivalente a maximizar el valor de $1 - \hat{y}$.
- Por el otro lado, el objetivo de la red *G* es lograr que la *D* produzca salidas lo mayores posibles para los ejemplos falsos generados por la red *G*. En definitiva las redes *G* y *D* se deben entrenar con objetivos opuestos.



1. Redes Generativas

- En otras palabras, este tipo de redes plantean un juego Min-Max, en el que buscamos el conjunto óptimo de parámetros que satisfacen una minimización y una maximización al mismo tiempo de una función objetivo V donde las variables que podemos modificar son los pesos de la red G y D.

- En general las GANs pueden producir resultados muy realistas para muchos problemas

- Sin embargo tienen problemas de estabilidad y son difíciles de entrenar
  - "Mode Collapse": el generador se centra en producir un limitado conjunto de patrones que funcionan bien en "engañar" al discriminante. Se puede mejorar aumentando el tamaño de los batches, cambiando o regularizando el discriminante, y usando otros métodos de optimización.
  - Baja variabilidad y diversidad de los outputs
  - Gradientes bajos

2. Aprendizaje por refuerzo

- El Aprendizaje por refuerzo o Reinforcement Learning es una rama del aprendizaje automático que se centra en la toma de decisiones para **maximizar las recompensas acumuladas** en una situación determinada a través de la experiencia.

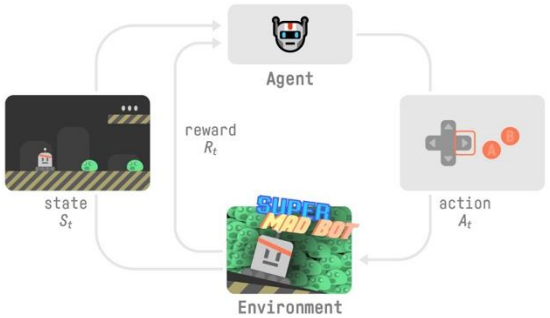
- En el RL, un **agente** aprende a lograr un objetivo en un entorno incierto y potencialmente complejo **realizando acciones y recibiendo retroalimentación a través de recompensas o penalizaciones**.

- Conceptos clave del aprendizaje por refuerzo
  - **Agente**: El aprendiz o tomador de decisiones.
  - **Entorno**: Todo aquello con lo que interactúa el agente.
  - **Estado**: Situación específica en la que se encuentra el agente.
  - **Acción**: Todos los movimientos posibles que puede realizar el agente.
  - **Recompensa**: Retroalimentación del entorno en función de la acción realizada.

2. Aprendizaje por refuerzo

- El proceso es el siguiente: el agente realiza acciones dentro del entorno, recibe recompensas o penalizaciones y ajusta su comportamiento para maximizar la recompensa acumulada.

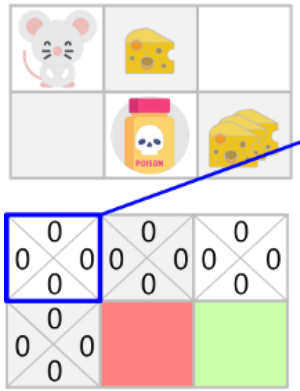
- Este proceso de aprendizaje se caracteriza por los siguientes elementos:
  - Política: Una estrategia utilizada por el agente para determinar la siguiente acción en función del estado actual.
  - Función de recompensa: una función que proporciona una señal de retroalimentación escalar basada en el estado y la acción.
  - Función de valor: una función que estima la recompensa acumulada esperada de un estado determinado.
  - Modelo del entorno: Una representación del entorno que ayuda en la planificación al predecir estados y recompensas futuras.



2. Aprendizaje por refuerzo

- Uno de los métodos basados en valor más habituales es el Q-learning, un algoritmo que entrena una función acción-valor

- La Q-function está resumida en una Q-table, una tabla en la que cada valor corresponde a un par de valor estado-acción. A medida que el agente explora el entorno la Q-tabla se actualiza.



States

Actions

	←	→	↑	↓
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0

2. Aprendizaje por refuerzo

- El algoritmo Q-Learning sigue los siguientes pasos:
  - Inicializamos la tabla con todos los valores igual a 0.
  - Escogemos una acción según un algoritmo exploitation-exploration: acción con mayor valor estado-acción o acción aleatoria
  - Actualizar la tabla con el siguiente valor:

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha[R_{t+1} + \gamma max_a Q(S_{t+1}, a) - Q(S_t, A_t)]$$

New Q-value estimation

Former Q-value estimation

Learning Rate

Immediate Reward

Discounted Estimate optimal Q-value of next state

Former Q-value estimation

TD Target

TD Error

- Al ser un método tabular no es escalable. En los últimos años se ha cambiado el algoritmo Q-Learning por Deep Q-Learning, en el que una red neuronal aproxima, dado un estado, los posibles Q-valores que cada posible acción puede dar en un estado.

3. Sistemas de recomendación

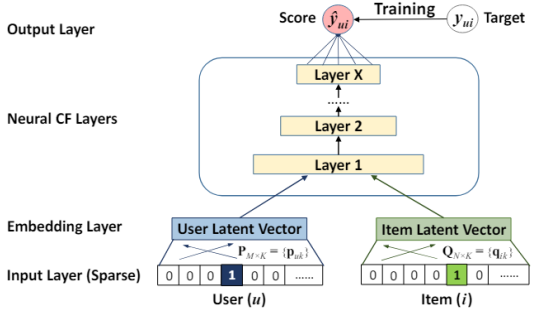
- Un sistema de recomendación es un algoritmo que da sugerencias de elementos que son más pertinentes para un usuario.

- Hay varios tipos de sistemas de recomendación, entre los que están:
  - Sistemas basados en contenido: usan las características de los artículos para la generación de las recomendaciones a través de la agrupación de productos. La técnica más habitual es la vectorización de las características de los productos y la utilización de distancias entre vectores (por ejemplo, si un usuario está visualizando una camiseta roja, el sistema de recomendación basado en contenido le recomendaría una selección de prendas parecidas o “cercanas”: otras camisetas rojas, etc)
  - Filtros colaborativos: realizan las recomendaciones al observar las preferencias de un usuario y compararlas con las de otros usuarios que tienen gustos similares. La técnica más habitual es el uso de modelos de machine learning (por ejemplo, si un usuario está visualizando una camiseta roja, el filtro colaborativo buscaría qué prendas han comprado usuarios que visualizaron la misma camiseta y se los enseñaría)

3. Sistemas de recomendación

- Uno de los modelos más populares se basa en redes neuronales. El Filtrado Neuronal Colaborativo (NCF) es una red neuronal que tiene por objetivo predecir la preferencia de un individuo por un producto, donde el par usuario- ítem es la entrada al algoritmo y la salida la predicción realizada en términos de probabilidad.

- La arquitectura del NCF tiene las siguientes partes:
  - La capa de embeddings de usuario y artículo para capturar las características de ambos.
  - Generalized Matrix Factorization (GMF): detecta relaciones lineales al multiplicar los embeddings de usuario y artículo.
  - Multi-Layer Perceptron (MLP): captura patrones no lineales a partir de los embeddings.
  - NeuMF: combina los resultados de GMF y MLP con una capa oculta y una función de activación Sigmoide.

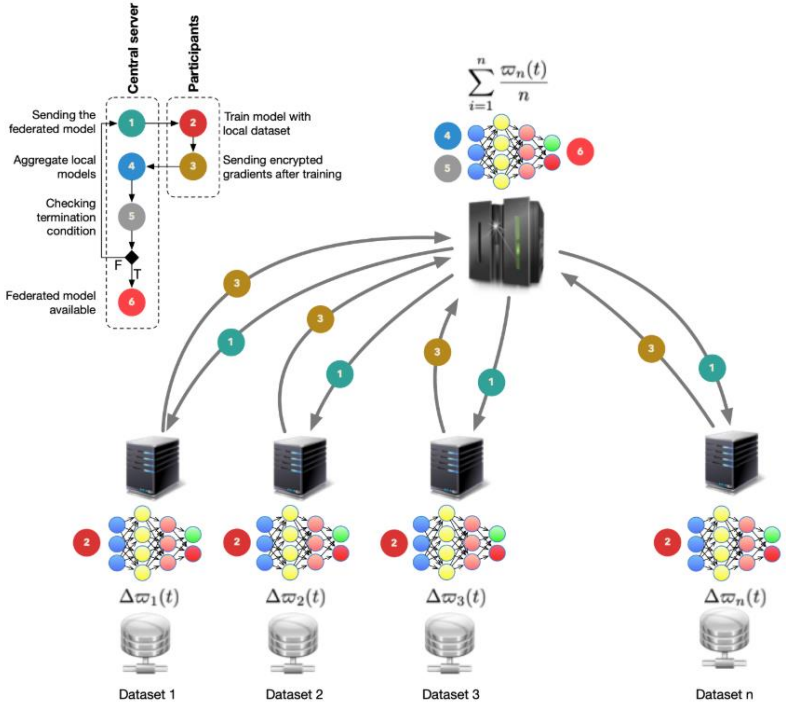


4. Federated Learning

- Federated learning o aprendizaje federado es un nuevo enfoque de aprendizaje automático distribuido que permite a varios participantes entrenar colaborativamente un modelo de aprendizaje automático de una manera privada y segura, y más importante, sin compartir datos privados entre los participantes o con un servidor centrar.

- El proceso inicial para entrenar un modelo de aprendizaje está formado por un servidor central y varios participantes con datos privados que quieren entrenar una red neuronal común. Los pasos son los siguientes:
  - El servidor central define la arquitectura de la red neuronal que será entrenada por todos los participantes y se la envía.
  - Los participantes entrenan el modelo usando sus datos privados, con lo que obtienen un modelo local, y después envían los parámetros (pesos o gradientes de la red) al servidor.
  - El servidor central agrega los parámetros del modelo usando una media aritmética. Con este proceso el servidor construye un modelo federado usando esos parámetros agregados.
  - El servidor envía el modelo federado de vuelta a los participantes para que lo reentrenen en sus datos locales y se itera el proceso, en general varios cientos de iteraciones.

4. Federated Learning



Redes Neuronales y Aprendizaje profundo

Tema 4 – Redes Generativas y Estado del Arte

Irina Arévalo